

[On page 22, please replace the second paragraph with the following:]

If a cDNA is identified as substantially identical to a known sequence in a public database, it is assigned an identifier which is the name and the accession number of the sequence with which it is substantially identical. In the case of a cDNA which represents the transcript of

2
3
cont
a human gene, it is also assigned a UniGene number (located on the world wide web at ncbi.nlm.nih.gov/UniGene and August 1996 NCBI News) if one is available. cDNAs which comprise subsequences which have substantial identity to one or more EST sequences in public databases are also assigned an EST number. cDNAs not having substantial identity to a sequence in a public database, are assigned an identifier designating the sequence as unknown and which is correlated in an array database with all available data relating to the sequence (e.g., sequence information, expression pattern, putative open reading frames, and motifs). In one embodiment of the invention, the user is provided with access to the array database when the user obtains the array.

[On page 22, please replace the third paragraph with the following:]

Search tools also include the Basic Local Alignment Search Tool 2 ("BLAST 2") used to align two given sequences and thereby identify regions having substantial sequence identity. Software for performing BLAST 2 analyses is publicly available through the National Center for Biotechnology Information (located on the world wide web at ncbi.nlm.nih.gov/). The BLAST algorithm performs a statistical analysis of the similarity between the two sequences provided (Tatiana A. Tatusova, Thomas L. Madden (1999), "Blast 2 sequences - a new tool for comparing protein and nucleotide
